Stock Price Prediction Model

Name:	**Saragadam Abhiram Laxmi Raj**
Registration No./Roll No.:	19274
Institute/University Name:	IISER Bhopal
Program/Stream:	EECS
Problem Release date:	February 02, 2022
Date of Submission:	April 24, 2022

Introduction

The stock price of a company is a great tool to analyze it's performance. The stock price may increase or decrease depending on the decisions made by the company or investor interests. Since it is highly volatile, proper analysis is of paramount importance to understand the risks involved before making an investment.

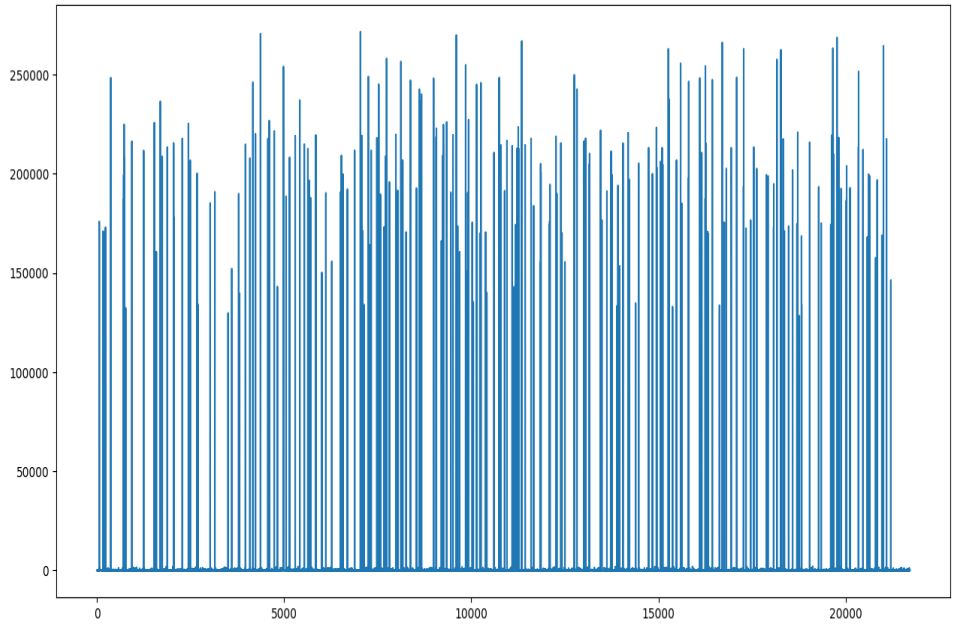

Figure 1: Test Data

1 Methods

We have used **linear regression**, **decision trees regression** and **SVR** to develop a model which can predict closing prices of various stock prices.

1.1 Getting our data

- Imports: Importing necessary libraries such as pandas, numpy, seaborn, matplotlib, sklearn etc.,
- Fixing our data: After uploading the data into our coding environment, we found three missing values in each column and we removed those data points as they gave better performance when compared to replacing the data with the mean/median.

1.2 Linear Regression

- Importing LinearRegression, metrics, mean-squared-error, r2-score and seaborn from sklearn.
- Splitting the data into training and testing sets.
- Fitting the training data into the linear regression model.
- Printing the prediction results based on the trained model.

1.3 Support Vector Regression

* Importing SVR, StandardScaler, MinMaxScaler, GridSearchCV from sklearn.svm, sklearn.preproces sklearn.modelselection respectively.
* Transformation of features by scaling each feature of the data.
* Fitting our data: Splitting the data into training and testing data and trained the model using testing data.
* The performance of the model is then improved by using parameter tuning, setting kernel = 'rbf', C = 1e3, gamma = 0.1
* The closing prices are then predicted and Mean squared error is calculated.

1.4 Decision Tree Regression

- Importing DecisionTreeRegressor, GridSearchCV
- Splitting the data into training and testing sets.
- Fitting the training data into the linear regression model.
- Printing the prediction results based on the trained model.

2 Evaluation Criteria

We evaluated the effectiveness of our model based on **MAE** and the **R-Squared Score**.

2.1 Evaluation Metrics

- Mean Absolute Error (MAE) is a measure of errors between paired observations expressing the same phenomenon. It is thus an arithmetic average of the absolute errors.
- R-Squared Score is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model.

3 Analysis of Results

From the three methods which we evaluated above, we got the best performance results using Linear Regression, when compared to Decision Trees and Support Vector Regression. We arrived on this conclusion based on the evaluation metric scores.

The MAE for Linear Regression is 4.6819.
The R-Squared Score for Linear Regression is 0.99999.

The MAE for Decision Tree Regression is 10.23.
The R-Squared Score for Decision Tree Regression is 0.99996.

The MAE for Support Vector Regression is 0.089.

4 Discussions and Conclusion

With the observations made in the model, we can conclude that the Linear Regression Model is an effective way to predict values for a regression problem, and can produce better results in comparison to other methods, in our case Decision Trees Regression and Support Vector Regression.

Here is the GitHub Link
https://github.com/sabhiram6/Stock_Price_Prediction